

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 08:43:50 ; Search time 168 Seconds  
(without alignments)  
657,501 Million cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380  
Sequence: 1 MAISIKTPEDIEKRVAGRL.....CELLLRKDDTIPAIISHDE 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubppaa/PCF\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubppaa/PCFUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*
  - 21: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
  - 22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	264	15	US-10-299-867-12
2	1380	100.0	264	15	US-10-350-516-8
3	1380	100.0	264	17	US-10-893-671-31
4	1380	100.0	264	18	US-10-813-549-1
5	1371	99.3	263	15	US-10-350-516-9
6	910.5	66.0	268	15	US-10-299-867-13
7	910.5	66.0	268	18	US-10-958-216-121
8	909.5	65.9	261	18	US-10-958-216-59
9	909.5	65.9	261	18	US-10-958-216-61
10	904.5	65.5	268	18	US-10-958-216-123
11	669.5	48.5	253	15	US-10-299-867-10

12	669.5	48.5	253	15	US-10-350-516-4	Sequence 4, Appl1
13	646	46.8	274	15	US-10-299-867-11	Sequence 11, Appl1
14	619	44.9	285	14	US-10-156-761-13458	Sequence 13458, A
15	617.5	44.7	331	16	US-10-425-115-299292	Sequence 299292, A
16	614	44.5	394	15	US-10-299-867-2	Sequence 2, Appl1
17	614	44.5	394	15	US-10-350-516-3	Sequence 3, Appl1
18	614	44.5	394	16	US-10-723-860-1103	Sequence 1103, Ap
19	614	44.5	444	15	US-10-399-645-13	Sequence 13, Appl1
20	595.5	43.2	263	9	US-09-738-626-5698	Sequence 5698, Ap
21	595	43.1	374	20	US-11-097-143-31929	Sequence 31929, A
22	590	42.8	248	18	US-10-510-812-34	Sequence 34, Appl1
23	587	42.5	337	15	US-10-424-599-247413	Sequence 247413, A
24	583	42.2	248	15	US-10-299-867-15	Sequence 15, Appl1
25	574.5	41.0	371	14	US-10-032-585-7897	Sequence 7897, Ap
26	565.5	41.0	348	16	US-10-425-115-185115	Sequence 185115, A
27	565	40.9	305	15	US-10-299-867-17	Sequence 17, Appl1
28	565	40.9	305	15	US-10-350-516-7	Sequence 7, Appl1
29	562.5	40.8	354	15	US-10-425-114-64761	Sequence 64761, A
30	559.5	40.5	367	15	US-10-299-867-9	Sequence 9, Appl1
31	557.5	40.4	346	16	US-10-437-963-150622	Sequence 150622, A
32	556.5	40.3	369	15	US-10-042-865-187	Sequence 187, App
33	552.5	40.0	369	15	US-10-042-865-187	Sequence 173546, A
34	550.5	39.9	356	15	US-10-424-599-173546	Sequence 173546, A
35	547.5	39.7	341	16	US-10-425-115-353435	Sequence 353435, A
36	543.5	39.4	373	16	US-10-437-963-170579	Sequence 170579, A
37	539	39.1	294	15	US-10-289-762-1081	Sequence 1081, Ap
38	531.5	38.5	382	15	US-10-424-599-244001	Sequence 244001, A
39	519	37.6	254	18	US-10-958-216-103	Sequence 103, App
40	518.5	37.6	307	15	US-10-042-865-188	Sequence 188, App
41	518.5	37.6	307	20	US-11-097-143-10797	Sequence 10797, A
42	518.5	37.6	350	16	US-10-425-115-299386	Sequence 299386, A
43	516	37.4	262	18	US-10-501-282-3480	Sequence 3480, Ap
44	515	37.3	254	18	US-10-958-216-105	Sequence 105, App
45	501	36.3	306	15	US-10-425-114-58168	Sequence 58168, A

# ALIGNMENTS

RESULT 1  
US-10-299-867-12  
Sequence 12, Application US/10299867  
GENERAL INFORMATION:  
APPLICANT: Symphon, Carolyn J.  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Frazier, Ronald B.  
APPLICANT: Woods, Cynthia L.  
APPLICANT: Zakeri, Hamideh  
TITLE OF INVENTION: Human methionine aminopeptidase type 3  
FILE REFERENCE: S03181-01-US  
CURRENT APPLICATION NUMBER: US/10/299,867  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: US 60/125,139  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 09/523,263  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 264  
TYPE: PRT  
ORGANISM: E. coli Metap

Query Match 100.0% ; Score 1380; DB 15; Length 264;  
Best Local Similarity 100.0% ; Pred. No. 2,2e+13;  
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAISIKTPEDIEKRVAGRLAEVLEMIPEYKPGVSTGELDRICNDYIVNEQHAVSACL 60



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 3, 2005, 08:30:15 ; Search time 165 Seconds

(without alignments)  
618.817 Million cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380  
Sequence: 1 MAISIKTPEDIKRMVAGRL.....CEILTKKDTTIPATISHDE 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003s:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	264	1	AAP70096 Met-amino
2	1380	100.0	264	4	AAU29345 Novel mar
3	1380	100.0	264	7	ADG44839 E.coli me
4	1380	100.0	264	7	ADG25269 E.coli m
5	1380	100.0	264	7	ADG42853 E.coli m
6	1371	99.3	263	7	ADG25270 E.coli m
7	1330	96.4	281	7	ABO65872 Klebsiell
8	1337.5	82.4	266	6	ABM67554 Photorhab
9	1086	78.7	268	7	ADP06311 Bacterial
10	910.5	66.0	268	7	ADG44840 H. influe
11	910.5	66.0	268	7	ADG42854 H. influe
12	910.5	66.0	268	7	ADG73413 Haemophil
13	909.5	65.9	261	7	ADG73351 P aerugin
14	909.5	65.9	285	7	ADG73353 P aerugin
15	909.5	65.9	285	7	ABO71996 Pseudomon
16	904.5	65.5	268	7	ADG73415 Haemophil
17	824.5	59.7	286	6	ADG33147 Acinetoba
18	799.5	57.9	268	8	ADU04891 M. catarr
19	798	57.8	264	6	ADA36289 Acinetoba
20	739	53.6	303	7	ABO82525 Pseudomon
21	701	50.8	265	7	ADP07562 Bacterial
22	682	49.4	264	2	AAW89276 Granulocy
23	669.5	48.5	253	7	ADG44837 Synechocy
24	669.5	48.5	253	7	ADG5285 Synechocy
25	669.5	48.5	253	7	ADG42851 Yeast met

26	646	46.8	274	7	ADG44838 Synechocy
27	646	46.8	274	7	ADG42852 Synechocy
28	616.5	44.7	290	3	AAQ35219 Zea mays
29	616.5	44.7	390	3	AAQ35218 Zea mays
30	614.5	44.5	291	3	AAQ33401 Zea mays
31	614.5	44.5	391	3	AAQ33400 Zea mays
32	614	44.5	394	7	ADG44829 Human met
33	614	44.5	394	7	ADG42864 Human met
34	614	44.5	394	7	ADG42843 Human met
35	614	44.5	394	8	ADQ18285 Human sof
36	614	44.5	494	5	AAU98895 Human pro
37	612.5	44.4	215	6	ABP76973 N. gonorr
38	595.5	43.2	263	4	AAQ31944 C. glutam
39	595.5	43.2	291	4	AAQ35219 Zea mays
40	595	43.1	374	4	ABQ68379 Drosophi
41	594	43.0	398	3	AAQ40455 Arabidops
42	594	43.0	398	3	AAQ20001 Arabidops
43	594	43.0	398	8	ADR42905 Arabidops
44	590	42.8	248	7	ADG92488 B. lichen
45	583	42.2	248	7	ADG44842 B. subtil

## ALIGNMENTS

RESULT 1  
AAP70096  
ID AAP70096 standard; protein; 264 AA.  
XX  
AC AAP70096;  
XX  
DT 09-APR-1991 (first entry)  
XX  
DB Met-aminopeptidase.  
XX  
KM Met-aminopeptidase; antibody; plasmid pSC1174.  
XX  
OS Escherichia coli.  
XX  
PN EP219237-A.  
PD 22-APR-1987.  
XX  
PF 19-SEP-1986; 86EP-00307242.  
XX  
PR 20-SEP-1985; 85US-00778414.  
PR 06-MAY-1986; 86US-00860330.  
XX  
PA (CERTU) CERTUS CORP.  
XX  
PI Benbasat A, Bauer KA, Chang S, Chang SY;  
XX WPI; 1987-110172/16.  
XX N-PSDB; AAN70151.  
XX N-terminal methionine free proteins prodn. - by using host transformed  
PT with vector to express a methionine-amino-peptidase.  
XX  
PS Disclosure; Fig 2; 20pp; English.  
XX  
CC The enzyme (from E.coli) is encoded by the 1.2 kb insert of pSC1174.  
CC proteins, esp. foreign proteins, lacking an N-terminal Met can be  
CC produced in bacterial systems, esp. interleukin-2 or ricin A. Antibodies  
CC to the enzyme are used for identifying such enzymes from various  
CC microorganisms. The enzyme pred. hydrolyses the N-terminal Met from Met-Y  
CC -Z (Y=Ala, Pro, Ile or Gly, Z=amino acid) and does not hydrolyse  
CC dipeptides  
XX  
SQ Sequence 264 AA;  
Query Match 100.0%; Score 1380; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 3e-140;  
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OM protein - protein search, using sw model

Run on: November 3, 2005, 08:31:04 ; Search time 175 Seconds  
(without alignments)  
772.508 Million cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380  
Sequence: 1 MAISIKTPEDIEKRVAGRL.....CELLTKRDKPTPAISHDE 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_crembl:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1380	100.0	264	1 AMPM_ECOLI	P07906 escherichia
2	1253	90.8	263	1 AMPM_SALTY	P10882 salmonella
3	1198	86.8	263	2 O66778	O66718 yersinia pe
4	1195	86.6	263	2 O62H67	O62H67 yersinia pe
5	1158	83.9	264	2 O6DBE4	O6DBE4 erwinia car
6	1137.5	82.4	265	2 O7N8P8	O7N8P8 photorhabdu
7	946	68.6	267	2 O9CNH2	O9CNH2 pasteurilla
8	937	67.9	264	1 AMPM_BUCAI	P57324 buchera ap
9	934	67.7	268	2 O7VP53	O7VP53 haemophilus
10	919.5	66.6	260	2 O886P4	O886P4 pseudomonas
11	915	66.3	264	2 O7VRE7	O7VRE7 candidatus
12	914.5	66.0	260	2 O88MI1	O88MI1 pseudomonas
13	910.5	65.9	268	1 AMPM_HARIN	P44421 haemophilus
14	909.5	65.4	261	1 AMPM_BUCAP	O9HXY1 pseudomonas
15	902	64.2	270	2 O65S27	O65S27 buchera ap
16	886.5	64.2	270	2 O65S27	O65S27 manheimia
17	881	63.8	258	2 O83BV1	O83BV1 coxiella bu
18	880.5	63.8	275	2 O6LND3	O6LND3 photobacter
19	878	63.6	258	2 O8PMV9	O8PMV9 xanthomonas
20	874	63.3	258	2 O8PAU5	O8PAU5 xanthomonas
21	853.5	61.8	259	2 O7QMC0	O7QMC0 anopheles g
22	853	61.8	258	2 O87F52	O87F52 xylella fas
23	852	61.7	259	2 O9PH35	O9PH35 xylella fas
24	852	61.7	253	2 O8DZG1	O8DZG1 wiggleswort
25	849.5	61.6	266	1 AMPM_BUCBP	O89AP3 buchera ap
26	843.5	61.1	265	2 O8EGH7	O8EGH7 shewanella
27	823.5	59.7	280	2 O9KPV1	O9KPV1 vibrio chol
28	804	58.3	272	2 O82XJ5	O82XJ5 nitrosomon
29	798	57.8	271	2 O62UC3	O62UC3 burkholderi
30	798	57.8	271	2 O63TJ1	O63TJ1 burkholderi
31	796.5	57.7	261	2 O66E99	O66E99 yersinia ps

32	796.5	57.7	261	2 O8ZBR7	O8ZBR7 yersinia pe
33	795	57.6	265	2 O7NVE6	O7NVE6 chromobacte
34	791	57.3	275	2 O8XZJ2	O8XZJ2 ralsionia s
35	779.5	56.5	254	2 O6FA52	O6FA52 acinetobact
36	775.5	56.2	292	2 O87MD7	O87MD7 vibrio para
37	775	56.2	263	2 O883J5	O883J5 pseudomonas
38	765.5	55.5	292	2 O8DBG2	O8DBG2 vibrio vuln
39	763.5	55.3	292	2 O7MTF9	O7MTF9 acinetobact
40	761	55.1	265	2 O6FV1	O6FV1 acinetobact
41	758.5	55.0	259	2 O9JXD9	O9JXD9 neisseria m
42	758.5	55.0	263	2 O7NVK8	O7NVK8 chromobacte
43	746	54.1	260	2 O73JY4	O73JY4 woldbachia p
44	745.5	54.0	260	2 O73JY4	O73JY4 bordetella
45	743	53.8	273	2 O7VYD1	O7VYD1 bordetella

#### ALIGNMENTS

RESULT 1  
ID AMPM\_ECOLI STANDARD; PRT; 264 AA.  
AC P07906;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Methionine aminopeptidase (EC 3.4.11.18) (MAP) (Peptidase M).  
GN Name=map;  
GN OrderedLocustNames=b0168, c0203, z0178, ECG0170, SF0158, S0161;  
OS Escherichia coli,  
OS Escherichia coli O6,  
OS Shigella flexneri,  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli;  
RX MEDLINE=87109068; PubMed=3027045;  
RA Ben-Bassat A., Bauer K., Chang S.-Y., Myambo K., Boosman A., Chang S.;  
RT "Processing of the initiation methionine from proteins: properties of  
the Escherichia coli methionine aminopeptidase and its gene  
structure.";  
RN J. Bacteriol. 169:751-757(1987).  
[2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=KL2 / W3110;  
RX MEDLINE=94261430; PubMed=8202364;  
RA Fujita N., Mori H., Yura T., Ishihama A.;  
RT "Systematic sequencing of the Escherichia coli genome: analysis of the  
2.4-4.1 min (110,917-193,643 bp) region.";  
RN Nucleic Acids Res. 22:1637-1639(1994).  
[3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=KL2 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.U.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RN Science 277:1453-1474(1997).  
[4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli;  
RX Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,  
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
RA Davis R.W.;  
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.

OM protein - protein search, using sw model

Run on: November 3, 2005, 08:34:05 ; Search time 40 Seconds  
(without alignments)  
635,031 Million cell updates/sec

Title: US-10-813-549-1

Sequence: 1 MAISTKPEDIEKRRVAGRL.....CEITLRKODTTPAISHDE 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: pirl:1  
2: pirl:2  
3: pirl:3  
4: pirl:4

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	264	1 DPECM	methionyl aminopep
2	1380	100.0	264	2 B85501	methionine aminope
3	1380	100.0	264	2 B90650	methionine aminope
4	1258	91.2	264	2 A10528	methionine aminope
5	1258	91.2	264	2 S12027	methionyl aminope
6	1195	86.6	263	2 AB0128	methionyl aminope
7	937	67.9	264	2 A64957	methionyl aminope
8	910.5	66.0	268	2 C64138	methionyl aminope
9	909.5	65.9	261	2 B83189	methionine aminope
10	852	61.7	259	2 B82845	methionine aminope
11	823.5	59.7	280	2 D82097	methionine aminope
12	796.5	57.7	261	2 AB0405	methionyl aminope
13	758.5	55.0	259	2 E82029	methionyl aminope
14	758.5	55.0	259	2 C81008	methionine aminope
15	739	53.6	260	2 C83301	probable methionin
16	723.5	52.4	256	2 AG2331	methionine aminope
17	687	50.5	276	2 H87380	methionine aminope
18	685.5	49.7	259	2 A71644	methionine aminope
19	684.5	49.6	259	2 D97859	methionyl aminope
20	669.5	48.5	253	2 S75731	methionyl aminope
21	660	47.8	278	2 AB2774	methionyl aminope
22	660	47.8	278	2 H97553	methionine aminope
23	659	47.8	276	2 A13341	methionyl aminope
24	646	46.8	274	2 S76999	methionyl aminope
25	634	45.9	285	2 B87106	methionine aminope
26	630	45.7	285	2 T50575	methionyl aminope
27	617	44.7	285	2 G70885	probable map prote
28	594	43.0	398	2 B84888	probable methionin
29	592	42.9	259	2 C75410	methionine aminope

30	583	42.2	248	2 J50493	methionyl aminope
31	568	41.2	275	2 AH1933	methionine aminope
32	565	40.9	305	2 S75743	methionyl aminope
33	559.5	40.5	387	2 S59390	methionyl aminope
34	557.5	40.4	250	2 F72247	methionine aminope
35	556.5	40.3	369	2 C86267	probable methionin
36	553.5	40.1	305	2 E85437	methionyl aminope
37	549	39.8	255	2 G81724	methionine aminope
38	543	39.3	291	2 D71462	probable methionin
39	540	39.1	248	2 T44405	methionyl aminope
40	539	39.1	266	2 A81531	methionine aminope
41	539	39.1	291	2 F86616	methionine aminope
42	539	39.1	291	2 E72008	methionine aminope
43	532	38.6	379	2 T40384	probable methionin
44	488.5	35.4	258	2 F70307	methionyl aminope
45	482	34.9	252	2 AD1660	methionine aminope

#### ALIGNMENTS

RESULT 1  
DPECM  
methionyl aminopeptidase (EC 3.4.11.18) [validated] - Escherichia coli (strain K-12)  
N/Alternate names: metallooligopeptidase; methionine aminopeptidase; peptidase M  
C/Species: Escherichia coli  
C/Date: 30-Sep-1988 #sequence revision 30-Sep-1988 #text\_change 09-Jul-2004  
R/Ben-Bassat, A.; Bauer, K.; Chang, S.Y.; Myambo, K.; Boosmaier, A.; Chang, S.  
J. Bacteriol. 169, 751-757, 1987  
A/Title: Processing of the initiation methionine from proteins: properties of the Esche  
A/Reference number: A27761; MUID:87109068; PMID:3027045  
A/Molecule type: DNA  
A/Residues: 1-264 <BEN>  
A/Cross-references: UNIPROT:P07906; GB:M15106; NID:9146726; PIDN:AAA24112.1; PID:914672  
A/Note: Comparative analyses indicate that residues adjacent to initial methionines of  
R.Fujita, N.  
submitted to the EMBL Data Library, January 1994  
A/Reference number: S45181  
A/Accession: S45233  
A/Molecule type: DNA  
A/Residues: 1-264 <FUJ>  
A/Cross-references: EMBL:D26562; NID:9473770; PIDN:BA05612.1; PID:9473823  
A/Experimental source: strain K-12, substrain W3110  
R/Blaetner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Ci  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A/Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503  
A/Accession: H64740  
A/Molecule type: DNA  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Residues: 1-264 <BLAT>  
A/Cross-references: GB:AE00126; GB:U00096; NID:91786358; PIDN:AACT3279.1; PID:91786364;  
A/Experimental source: strain K-12, substrain MG1655  
R/Roderick, S.L.; Matthews, B.W.  
submitted to the Brookhaven Protein Data Bank, December 1992  
A/Reference number: A51847; PDB:1MAT  
A/Contents: annotation; X-ray crystallography, 2.4 angstroms, residues 2-264  
R/Roderick, S.L.; Matthews, B.W.  
Biochemistry 32, 3907-3912, 1993  
A/Title: Structure of the cobalt-dependent methionine aminopeptidase from Escherichia co  
A/Reference number: A49470; MUID:93229467; PMID:8471602  
A/Contents: annotation; X-ray crystallography, 2.4 angstroms  
C/Genetics:  
A/Map position: 4 min  
C/Complex: monomer  
C/Function:  
A/Description: catalyzes hydrolysis of amino-terminal methionine from proteins  
C/Superfamily: Escherichia coli methionyl aminopeptidase  
C/Keywords: aminopeptidase; cobalt; metalloprotein; monomer; protein biosynthesis

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 11, 2005, 18:02:50 / Search time 2355 Seconds  
(without alignments)  
4267.077 Million cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380

Sequence: 1 MAISIKTPEDEIKKRVAGRL.....CEILTLRKDDTIPATISHDE 264

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=cgnt\_1/USPRO.spool/US10813549/runat\_03112005\_075935\_13175/app\_query.fasta\_1.455  
-DB=EST -QMT=fastlap -SUFFI=ret -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTFM=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US10813549 @CGN 1 1 3437 @runat\_03112005\_075935\_13175 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

Database: EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1351	97.9	812	9	CL661282 PRI0139b
2	1327	86.2	844	9	CL668009 PRI0148b
3	1236	89.6	716	9	CL663485 PRI0144b
4	793.5	57.5	830	1	AL669413 AL669413
5	786	57.0	706	8	AQ0989727 RfC00360
6	756	54.8	839	9	CL666394 PRI0152b
7	614	44.5	1047	9	AY420765 Homo sapi
8	607	44.0	1335	9	CR696222 Tetradon
9	605	43.8	1047	9	AY420767 Mus muscu

10	605	43.8	2550	3	AK077694	AK077694 Mus muscu
11	604	43.7	1860	3	CR694735	CR694735 Tetradon
12	603	43.7	870	7	CO163928	CO163928 FLD1_44_E
13	601	43.6	739	9	CL685899	CL685899 PRI0142b
14	599	43.4	631	8	BH762604	BH762604 BMBAC330A
15	599	43.4	995	9	AY420766	AY420766 Pan trogl
16	594	43.0	1308	3	CR688406	CR688406 Tetradon
17	585	42.4	926	5	BO706230	BO706230 AGENCOURT
18	578	41.9	890	7	CK458575	CK458575 AGENCOURT
19	573	41.5	859	5	BQ232415	BQ232415 AGENCOURT
20	569.5	41.3	967	3	AK006484	AK006484 Mus muscu
21	568	41.2	940	5	BO937757	BO937757 AGENCOURT
22	561.5	40.7	997	7	CK264740	CK264740 EST710818
23	560.5	40.6	780	8	BZ573414	BZ573414 msh2_3093
24	558.5	40.5	867	7	CF205959	CF205959 R8909151
25	558.5	40.5	885	7	CF514702	CF514702 Cabud0005
26	556.5	40.3	1235	3	CNS04BLO	BX161624 Arabidops
27	549.5	39.8	876	5	BUI36049	BUI36049 603124836
28	548.5	39.7	1278	3	CNS0A2S4	BX627291 Arabidops
29	547	39.6	742	4	BU434305	BU434305 BU434305
30	547	39.6	760	4	BU432312	BU432312 BU432312
31	545	39.5	888	7	CN159313	CN159313 948372 MA
32	542.5	39.3	637	1	AJ722958	AJ722958 AJ722958
33	542.5	39.3	637	1	CO978868	CO978868 GM89003A1
34	541	39.2	682	6	CD717720	CD717720 VVB142C09
35	541	39.2	761	4	BU434187	BU434187 B2554793 pacet1-60
36	540.5	39.2	1347	8	B2554793	B2554793 pacet1-60
37	537.5	38.9	915	7	CK269489	CK269489 EST715567
38	533	38.6	778	7	CN359624	CN359624 170006000
39	532.5	38.6	898	7	CY243178	CY243178 WS02518.B
40	530.5	38.4	989	7	CO025828	CO025828 EST804212
41	528	38.3	715	4	BB468358	BB468358 BB468358
42	527.5	38.2	684	2	BB468358	BB468358 BB468358
43	527.5	38.2	797	2	BB902721	BB902721 601675531
44	526	38.1	722	6	CA239368	CA239368 SCBFFL507
45	526	38.1	731	1	AV881422	AV881422 AV881422

## ALIGNMENTS

RESULT 1  
LOCUS CL661282 812 bp DNA linear GSS 09-JUN-2004  
DEFINITION PRI0139b.G03 - PRI0139b.B21 (812) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.  
ACCESSION CL661282  
VERSION CL661282.1 GI:50147601  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Neodiplogasteridae; Pristionchus.  
1 (bases 1 to 812)  
Strinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
Appabdi: an Acedb database for the nematode satelite organism  
Pristionchus pacificus  
Nucleic Acids Res 32 (1), D421-D422 (2004)  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.  
Seq primer: T7  
Class: fosmid ends.  
Location/Qualifiers  
1..812  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"

## FEATURES

source

OM protein - nucleic search, using frame\_plus\_g2n model

Run on: November 11, 2005, 16:26:32 ; Search time 3123 Seconds

(Without alignments)  
4096.118 Million cell updates/sec

Title: US-10-813-549-1  
Perfect score: 1380  
Sequence: 1 MAISIKTPEDIKRVAAGRL.....CELLLRKDDTTPAISHDE 264

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.epool/US10813549/runat.03112005.075935.13164/app.query.fasta\_1.455  
-DB=genEmb1 -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10813549 @cgn 1 1 3731 @runat.03112005.075935.13164 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESOURCE -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_on:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	1197	1	ECOMAP
2	1380	100.0	1197	6	E01194
3	1380	100.0	1197	6	E01194
4	1380	100.0	1197	6	I02543

C	5	1380	100.0	11204	6	AX370213	AX370213 Sequence
C	6	1380	100.0	13945	1	AE015052	AE015052 Shigella
C	7	1380	100.0	14042	1	AE005192	AE005192 Escherichia
C	8	1380	100.0	82727	1	EC082K	D26562 Escherichia
C	9	1380	100.0	110000	1	U00096_01	Continuation (2 of
C	10	1380	100.0	123171	1	ECU70214	U70214 Escherichia
C	11	1380	100.0	281530	1	AP002550	AP002550 Escherich
C	12	1380	100.0	290029	1	AE016978	AE016978 Shigella
C	13	1380	100.0	300409	1	AE016755	AE016755 Escherich
C	14	1380	100.0	303656	2	AC025528	AC025528 Mus muscu
C	15	1330	96.4	846	6	AR388489	AR388489 Sequence
C	16	1325	96.0	3779	1	ECOLNNA	M96431 Escherichia
C	17	1258	91.2	1811	1	STPEPMG	X55778 S. typhimur
C	18	1258	91.2	19971	1	AE008704	AE008704 Salmonell
C	19	1258	91.2	251050	1	AL627265	AL627265 Salmonell
C	20	1258	91.2	300169	1	AE016834	AE016834 Salmonell
C	21	1198	86.8	110000	1	BX36398_35	Continuation (36 o
C	22	1195	86.6	10443	1	AE013914	AE013914 Yersinia
C	23	1195	86.6	210050	1	AE014146	AE014146 Yersinia
C	24	1195	86.6	291817	1	AE017137	AE017137 Yersinia
C	25	1158	83.9	110000	1	BX950851_11	Continuation (12 o
C	26	1137.5	82.4	342905	1	BX571861	BX571861 Phototrab
C	27	1137.5	82.4	349980	6	AX770908	AX770908 Sequence
C	28	1086	78.7	807	6	AR377418	AR377418 Sequence
C	29	946	68.6	11341	1	AE006081	AE006081 Pasteurel
C	30	937	67.9	110000	6	AR409405_2	Continuation (3 of
C	31	937	67.9	110000	6	BD061520_2	Continuation (3 of
C	32	937	67.9	347550	1	AP001118	AP001118 Buchnera
C	33	934	67.7	304558	1	AE017151	AE017151 Haemophil
C	34	919.5	66.6	310029	1	AE016861	AE016861 Pseudom
C	35	915	66.3	267050	1	BX248585	BX248585 Blochmann
C	36	914.5	66.3	301995	1	AE016779	AE016779 Pseudom
C	37	911.5	66.1	308766	6	CO873161	CO873161 Sequence
C	38	910.5	66.0	807	7	AX928193	Continuation (18 o
C	39	910.5	66.0	11782	1	U32845	Continuation (18 o
C	40	910.5	66.0	110000	6	BD426631_17	Continuation (18 o
C	41	910.5	66.0	110000	6	AR274513_17	Continuation (18 o
C	42	910.5	66.0	110000	6	AR541453_17	Continuation (18 o
C	43	909.5	65.9	786	6	AX928131	AX928131 Sequence
C	44	909.5	65.9	786	6	AX928133	AX928133 Sequence
C	45	909.5	65.9	14537	1	AE004785	AE004785 Pseudom

ALIGNMENTS

RESULT 1	ECOMAP	1197 bp	DNA	linear	BCI 26-APR-1993
LOCUS	E.coli map gene, encoding methionine amino peptidase, complete cds.				
DEFINITION	E.coli map gene, encoding methionine amino peptidase, complete cds.				
ACCESSION	M5106				
VERSION	M5106.1 GI:146726				
KEYWORDS	map gene, metallo-oligopeptidase, methionine aminopeptidase.				
SOURCE	Escherichia coli				
ORGANISM	Escherichia coli				
REFERENCE	Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.				
AUTHORS	1 (bases 1 to 1197) Ben-Bassat,A., Bauer,K., Chang,S.Y., Myambo,K., Boosman,A. and Chang,S.				
TITLE	Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure				
JOURNAL	J. Bacteriol. 169 (2), 751-757 (1987)				
MEDLINE	87109068				
PUBMED	3027045				
COMMENT	Original source text: E.coli (strain CM89) DNA, clone PSYC1174. Draft entry and computer-readable sequence for [1] kindly provided by S.Chang, 01-JUN-1987. There are at least two tandem 'map' gene promoters separated by the SnaI site at positions 150. Two regions of dyad symmetry are located at positions 187-201 and 1018-1037.				
FEATURES	Location/Qualifiers				
source	1..1197				



OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 11, 2005, 18:29:14 ; Search time 621 Seconds

(without alignments)  
3515,672 Million cell updates/sec

Title: US-10-813-549-1  
Perfect score: 1380  
Sequence: 1 MA5IKTPEDEIKRVAAGRL.....CEITLRKDDTPIAISHDE 264

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 413490567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10813549/runat\_03112005\_075936\_13214/app\_query.fasta\_1.455  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rpmb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10813549 @cgn 1 1 480 @runat\_03112005\_075936\_13214  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGESUBSTR -NEG\_SCORES=0 -WAIT -DSHLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
23: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
25: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
26: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
27: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
28: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	11204	22 US-10-893-671-30	Sequence 30, Appl
2	937	67.9	640681	9 US-09-790-988-1	Sequence 1, Appl
3	911.5	66.1	908766	24 US-10-958-159-685	Sequence 665, Appl
4	910.5	66.0	807	24 US-10-958-216-120	Sequence 120, Appl
5	910.5	66.0	1830121	18 US-10-329-670-1	Sequence 1, Appl
6	910.5	66.0	1830121	21 US-10-158-865-1	Sequence 1, Appl
7	910.5	66.0	1830121	24 US-10-981-687-1	Sequence 1, Appl
8	909.5	65.9	786	24 US-10-958-216-58	Sequence 58, Appl
9	909.5	65.9	786	24 US-10-958-216-60	Sequence 60, Appl
10	904.5	65.5	807	24 US-10-958-216-122	Sequence 122, Appl
11	852	61.7	2731748	20 US-10-297-465A-1	Sequence 1, Appl
12	799.5	57.9	96109	19 US-10-672-787-35	Sequence 35, Appl
13	758.5	55.0	92934	24 US-10-915-740A-21	Sequence 21, Appl
14	758.5	55.0	2242716	24 US-10-915-740A-1068	Sequence 1068, Appl
15	619	44.9	855	16 US-10-156-761-5908	Sequence 5908, Appl
16	619	44.9	9025608	16 US-10-156-761-1	Sequence 1, Appl
17	617.5	44.7	1735	21 US-10-425-115-11629	Sequence 11629, Appl
18	614	44.5	2671	18 US-10-299-867-1	Sequence 1, Appl
19	614	44.5	2671	21 US-10-723-860-1102	Sequence 1102, Appl
20	614	44.5	2984	18 US-10-399-645-28	Sequence 28, Appl
21	597.5	43.3	3309400	9 US-09-738-626-1	Sequence 1, Appl
22	595.5	43.2	789	9 US-09-738-626-2198	Sequence 2198, Appl
23	595	43.1	1125	26 US-11-097-143-31928	Sequence 31928, Appl
24	594	43.0	1197	9 US-09-938-842A-2427	Sequence 2427, Appl
25	594	43.0	1197	11 US-09-938-842A-2427	Sequence 2427, Appl
26	590	42.8	747	9 US-09-974-300-1195	Sequence 1195, Appl
27	590	42.8	1747	24 US-10-510-812-33	Sequence 33, Appl
28	587	42.5	1605	19 US-10-424-599-104571	Sequence 104571, Appl
29	574.5	41.6	1165	17 US-11-032-586-6897	Sequence 6897, Appl
30	571.5	41.4	2287	26 US-11-097-143-31924	Sequence 31924, Appl
31	571.5	41.4	3252	26 US-11-097-143-31927	Sequence 31927, Appl
32	565.5	41.0	1360	21 US-10-425-115-452	Sequence 452, Appl
33	562.5	40.8	1295	19 US-10-425-115-2639	Sequence 2639, Appl
34	557.5	40.4	1306	20 US-10-437-963-48139	Sequence 48139, Appl
35	552.5	40.0	3088	21 US-10-723-860-5641	Sequence 5641, Appl
36	550.5	39.9	1534	19 US-10-424-599-30704	Sequence 30704, Appl
37	547.5	39.7	1889	21 US-10-425-115-168772	Sequence 168772, Appl
38	547.5	39.7	1489	20 US-10-437-963-68096	Sequence 68096, Appl
39	539	39.1	1230025	18 US-10-289-762-1	Sequence 1, Appl
40	531.5	38.5	1525	19 US-10-424-599-101159	Sequence 101159, Appl
41	519	37.6	762	24 US-10-958-216-102	Sequence 102, Appl
42	519	37.6	1856	9 US-09-070-927A-494	Sequence 494, Appl
43	518.5	37.6	943	26 US-11-097-143-10796	Sequence 10796, Appl
44	518.5	37.6	1734	21 US-10-425-115-114723	Sequence 114723, Appl
45	516	37.4	807	24 US-10-501-282-3477	Sequence 3477, Appl

ALIGNMENTS

RESULT 1  
US-10-893-671-30/c  
; Sequence 30, Application US/10893671  
; Publication No. US20050064527A1  
; GENERAL INFORMATION:  
; APPLICANT: Levy, Stuart, et. al.  
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE  
; FILE REFERENCE: PKZ-043  
; CURRENT APPLICATION NUMBER: US/10/893,671  
; PRIOR FILING DATE: 2004-07-15  
; PRIOR APPLICATION NUMBER: US/09/801,563  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/188,362  
; PRIOR FILING DATE: 2000-03-10



OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 11, 2005, 18:19:03 ; Search time 145 Seconds

(without alignments)  
2979.152 Million cell updates/sec

Title: US-10-813-549-1  
Perfect score: 1380  
Sequence: 1 MAISIKTPEDERKRVAGRL.....CEITTKKDTTPAISHDE 264

Scoring table:  
BIOSUN62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-O=/cg2\_1/USPTO.spool/US10813549/runat\_03112005\_075936.13189/app.query.fasta\_1.455  
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum2 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR.SCORB=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=pct -NOM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10813549 -CGEN\_1.1\_69 @runat\_03112005\_075936.13189 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents NA:  
1: /cg2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cg2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cg2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cg2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cg2\_6/ptodata/1/ina/PTUS.COMB.seq.\*  
6: /cg2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1330	96.4	846	4	US-09-489-039A-5218 Sequence 5218, Ap
2	1086	78.7	807	4	US-09-543-681A-2424 Sequence 2424, Ap
3	937	67.9	640681	4	US-09-790-988-1 Sequence 1, Appl
4	910.5	66.0	1830121	4	US-09-557-884-1 Sequence 1, Appl
5	910.5	65.9	858	4	US-09-252-991A-4171 Sequence 4171, Ap
6	909.5	65.9	963	4	US-09-252-991A-4435 Sequence 4435, Ap
7	909.5	65.9	1176	4	US-09-252-991A-4314 Sequence 4314, Ap
8	909.5	65.9	861	4	US-09-328-352-308 Sequence 308, Ap
9	799.5	57.9	807	4	US-09-540-236-657 Sequence 657, Ap
10	799.5	57.9	96109	4	US-09-596-002-35 Sequence 35, Appl
11	799.5	57.8	795	4	US-09-328-352-3450 Sequence 3450, Ap
12	798	57.8	795	4	US-09-328-352-3450 Sequence 3450, Ap

13	739	53.6	912	4	US-09-252-991A-14700 Sequence 14700, A
14	739	53.6	1125	4	US-09-252-991A-14568 Sequence 14568, A
15	739	53.6	1248	4	US-09-252-991A-15057 Sequence 15057, A
16	701	50.8	798	4	US-09-543-681A-3675 Sequence 3675, Ap
17	682	49.4	2706	3	US-09-066-046-23 Sequence 23, Appl
18	617	44.7	852	4	US-09-902-540-4170 Sequence 4170, Ap
19	617	44.7	18686	4	US-09-902-540-4170 Sequence 1206, Ap
20	617	44.7	4403765	3	US-09-103-840A-2 Sequence 1, Appl
21	617	44.7	4411529	3	US-09-103-840A-2 Sequence 1, Appl
22	614	44.5	2671	4	US-09-523-263B-1 Sequence 1, Appl
23	597.5	43.3	966	4	US-09-602-777A-35 Sequence 35, Appl
24	575	41.7	1239	4	US-09-270-767-10002 Sequence 10002, A
25	552.5	40.0	783	4	US-09-902-540-5059 Sequence 5059, Ap
26	552.5	40.0	30135	4	US-09-902-540-1249 Sequence 1249, Ap
27	539	39.1	123025	4	US-09-198-452A-1 Sequence 1, Appl
28	539	39.1	1230230	4	US-09-438-185A-1 Sequence 1, Appl
29	489.5	35.5	1529	4	US-09-523-263B-7 Sequence 7, Appl
30	488.5	35.5	1798	3	US-09-797-906-1 Sequence 1, Appl
31	489.5	35.5	3064	4	US-09-620-312D-378 Sequence 378, App
32	489.5	35.5	3114	4	US-09-523-263B-5 Sequence 5, Appl
33	473	34.3	618	4	US-09-489-039A-2229 Sequence 2229, Ap
34	450.5	32.6	1039	3	US-08-975-762-16 Sequence 36, Appl
35	450.5	32.6	1039	3	US-08-821-324-36 Sequence 36, Appl
36	450.5	32.6	1039	3	US-09-235-028-36 Sequence 36, Appl
37	450.5	32.6	1039	3	US-09-106-582-36 Sequence 36, Appl
38	450.5	32.6	1039	4	US-09-159-469-36 Sequence 36, Appl
39	450.5	32.6	1039	4	US-09-693-542-36 Sequence 36, Appl
40	450.5	32.6	3073	3	US-08-975-762-41 Sequence 41, Appl
41	450.5	32.6	3073	3	US-09-235-028-41 Sequence 41, Appl
42	450.5	32.6	3073	3	US-09-106-582-41 Sequence 41, Appl
43	450.5	32.6	3073	4	US-09-159-469-41 Sequence 41, Appl
44	450.5	32.6	3073	4	US-09-693-542-41 Sequence 41, Appl
45	446.5	32.4	756	4	US-09-710-279-1895 Sequence 1895, Ap

# ALIGNMENTS

RESULT 1  
US-09-489-039A-5218  
Sequence 5218, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: 2709.2004001  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US/09/489,039A  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 5218  
LENGTH: 846  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5218

# Alignment Scores:

Pred. No.: 2,02e-174  
Score: 1330.00  
Percent Similarity: 97.73%  
Best Local Similarity: 96.21%  
Query Match: 96.38%  
DB: 4  
Length: 846  
Matches: 254  
Conservative: 4  
Mismatch: 6  
Indels: 0  
Gaps: 0

US-10-813-549-1 (1-264) x US-09-489-039A-5218 (1-846)  
QY 1 MetAlaISerIleThpProGluAaPIIleGluLyMetArgValaIaGlyArgLeu 20  
Db 52 ATGGCTATCTCTATTAGACATCTGAAACATCGAAATATGCGCTGCGCCGCTG 111  
QY 21 AAlaAGValleuGluMetIleGluProTyValIleProGlyValSerThryGlu 40

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 11, 2005, 14:27:17 ; Search time 406 Seconds

(without alignments)  
8449.290 Million cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380

Sequence: 1 MAISIKTPEDIEKRVAGRL.....CEILTLRKDDTIPALISHDE 264

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-O=/cgn2/1/USPTO/spool/US10813549/runat 03112005 075935 1156/app query.fasta.1.455  
-DB=N Geneseg 16Dec04 -OPMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0  
-LOOEXT=0 -INITTS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=PCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10813549 @CCGN 1.1 470 @runat 03112005 075935 1156 -NCPU=6 -ICPU=3  
-NO MAMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database:

N\_Geneseg\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	1197	1	AAAN70151
2	1380	100.0	11204	4	AAAS46245
3	1330	96.4	846	11	ACH99423
4	1137.5	82.4	798	10	ACF69733
5	1137.5	82.4	110000	10	ACF67367_27

C	6	1137.5	82.4	110000	10	ACF65386_1	Continuation (2 of
C	7	1086	78.7	807	10	ADFO2139	Adf02139 Bacterial
C	8	937	67.9	110000	6	ABA92787_2	Continuation (3 of
C	9	911.5	66.1	308766	13	ADT05738	Adt05738 Haemophil
C	10	910.5	66.0	807	10	ADG73412	Adg73412 Haemophil
C	11	910.5	66.0	110000	2	AA742063_17	Continuation (18 o
C	12	909.5	65.9	786	10	ADG73350	Adg73350 Pseudomon
C	13	909.5	65.9	786	10	ADG73352	Adg73352 Pseudomon
C	14	909.5	65.9	858	11	ABD05567	Abd05567 Pseudomon
C	15	909.5	65.9	963	11	ABD05831	Abd05831 Pseudomon
C	16	909.5	65.9	1176	11	ABD05710	Abd05710 Pseudomon
C	17	904.5	65.5	807	10	ADG73414	Adg73414 Haemophil
C	18	824.5	59.7	807	10	ADA29021	Ada29021 DNA encod
C	19	799.5	57.9	807	12	ADL02971	Adl02971 DNA encod
C	20	799.5	57.9	96109	4	AAF28548	AAF28548 Genomic f
C	21	798	57.8	795	9	ADA32163	Ada32163 DNA encod
C	22	758.5	55.0	92934	3	AAA81473	AAA81473 N. mening
C	23	758.5	55.0	110000	3	AAAB1489_7	Continuation (8 of
C	24	758.5	55.0	172325	3	AAF21613	AAF21613 Neisseria
C	25	739	53.6	912	11	ABD16096	Abd16096 Pseudomon
C	26	739	53.6	1125	11	ABD15964	Abd15964 Pseudomon
C	27	739	53.6	1248	11	ABD16453	Abd16453 Pseudomon
C	28	701	50.8	798	10	ADF03390	Adf03390 Bacterial
C	29	682	49.4	2706	2	AAV81790	AAV81790 Granulocy
C	30	618.5	44.8	645	10	AB237943	AB237943 N. gonorr
C	31	617	44.7	110000	4	AAI99682_31	Continuation (32 o
C	32	617	44.7	110000	4	AAI99682_31	Continuation (32 o
C	33	616.5	44.7	1625	3	AAAC4483	AAAC4483 Zea mays
C	34	614.5	44.5	1578	3	AAAC4373	AAAC4373 Zea mays
C	35	614	44.5	2671	10	ADG44828	Adg44828 Human CDN
C	36	614	44.5	2671	10	ADG42842	Adg42842 Human met
C	37	614	44.5	2671	12	AD018284	AD018284 Human sof
C	38	614	44.5	2984	6	ABK86146	ABK86146 CDNA enco
C	39	597.5	43.3	996	4	AAAF71155	AAAF71155 Corynebac
C	40	597.5	43.3	349980	5	AAH68531	AAH68531 C glutami
C	41	597.5	43.3	349980	5	AAH68530	AAH68530 C glutami
C	42	595.5	43.2	789	5	AAH67163	AAH67163 C glutami
C	43	595	43.1	1125	4	ABL23125	ABL23125 Drosophi
C	44	594	43.0	1197	6	ABZ14622	ABZ14622 Arabidops
C	45	594	43.0	1419	13	ADR42904	ADR42904 Arabidops

# ALIGNMENTS

RESULT 1	
ID	AAAN70151 standard; DNA; 1197 BP.
XX	
XX	AAAN70151;
XX	
DT	09-APR-1991 (first entry)
XX	
DE	Insert of psyc1174 encoding Met-aminopeptidase.
XX	
KM	Met-aminopeptidase; antibody.
OS	Escherichia coli.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	219..1013
XX	/*tag= a
PN	EP219237-A.
XX	
PD	22-APR-1987.
XX	
PF	19-SEP-1986; 86EP-00307242.
XX	
PR	20-SEP-1985; 85US-00778414.
XX	
PA	06-MAY-1986; 86US-00860330.
XX	
XX	(CETU ) CETUS CORP.